US006447773B1

(12) **United States Patent**
Black et al.

(10) **Patent No.: US 6,447,773 B1**
(45) **Date of Patent: Sep. 10, 2002**

(54) **MEMBER OF ATP DEPENDENT PROTEASE FAMILY, CLPL**

(75) Inventors: **Michael Terence Black**, Chester Springs; **John Edward Hodgson**, Malvern, both of PA (US); **David Justin Charles Knowles**, Redhill (GB); **Raymond Winfield Reichard**, Quakertown, PA (US); **Richard Oakley Nicholas**, Collegeville, PA (US); **Martin Karl Russel Burnham**, Norristown, PA (US); **Julie M Pratt**, Verona (IT); **Martin Rosenberg**, Royersford, PA (US); **Judith M Ward**, Dorking Surrey; **Andrew Fosberry**, Linton, both of (GB); **Elizabeth J Lawlor**, Malvern, PA (US)

(73) Assignees: **SmithKline Beecham Corporation**, Philadelphia, PA (US); **SmithKline Beecham plc**, Brentford (GB)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/416,874**

(22) Filed: **Oct. 13, 1999**

Related U.S. Application Data

(60) Division of application No. 08/923,511, filed on Sep. 4, 1997, now Pat. No. 6,274,376, which is a continuation-in-part of application No. PCT/US97/02318, filed on Feb. 19, 1997.

(60) Provisional application No. 60/011,888, filed on Feb. 20, 1996.

(51) **Int. Cl.**[7] **A61K 38/48**

(52) **U.S. Cl.** **424/94.63**; 530/350; 424/94.64; 514/2; 514/12; 435/195; 435/212; 435/219; 435/220

(58) **Field of Search** 530/350, 324; 435/195, 212, 219, 220; 424/94.63, 94.64; 514/2, 12

(56) **References Cited**

PUBLICATIONS

Dao Chao Huang et al., "Two genes present on a transposon–like structure in *Lactococcus lactis* are involved in a Clp–family proteolytic activity", Molecular Microbiology (1993) vol. 7, No. 6 pp. 957–965.*

Huang, et al., GenBank Submission, Accession No. X62333, "Two genes present on a transposon–like structure in *Lactococcus lactis* are involved in a Clp–family proteolytic activity."

* cited by examiner

Primary Examiner—Marianne P. Allen
(74) *Attorney, Agent, or Firm*—Jason C. Fedon; Edward R. Gimmi; Charles M. Kinzig

(57) **ABSTRACT**

The invention provides clpL polypeptides and DNA (RNA) encoding clpL polypeptides and methods for producing such polypeptides by recombinant techniques. Also provided are methods for utilizing polypeptides to screen for antibacterial compounds.

6 Claims, No Drawings

MEMBER OF ATP DEPENDENT PROTEASE FAMILY, CLPL

This application is a divisional application of U.S. patent application Ser. No. 08/923,511, filed Sep. 4, 1997, now U.S. Pat. No. 6,274,376, which is a continuation-in-part of International Patent Application PCT/US97/02318, filed Feb. 19, 1997, which claims benefit to U.S. Provisional Application Ser. No. 60/011,888, filed Feb. 20, 1996.

FIELD OF THE INVENTION

This invention relates to newly identified polynucleotides and polypeptides, and their production and uses, as well as their variants, agonists and antagonists, and their uses. In particular, in these and in other regards, the invention relates to novel polynucleotides and polypeptides of the ATP dependent proteases family, hereinafter referred to as "clpL".

BACKGROUND OF THE INVENTION

It is particularly preferred to employ Staphylococcal genes and gene products as targets for the development of antibiotics. The Staphylococci make up a medically important genera of microbes. They are known to produce two types of disease, invasive and toxigenic. Invasive infections are characterized generally by abscess formation effecting both skin surface and deep tissues. *Staphylococcus aureus* is the second leading cause of bacteremia in cancer patients. Osteomyelitis, septic arthritis, septic thrombophlebitis and acute bacterial endocarditis are also relatively common. There are at least three clinical conditions resulting from the toxigenic properties of Staphylococci. The manifestation of these diseases result from the actions of exotoxins as opposed to tissue invasion and bacteremia. These conditions include: Staphylococcal food poisoning, scalded skin syndrome and toxic shock syndrome.

The frequency of *Staphylococcus aureus* infections has risen dramatically in the past 20 years. This has been attributed to the emergence of multiply antibiotic resistant strains and an increasing population of people with weakened immune systems. It is no longer uncommon to isolate *Staphylococcus aureus* strains which are resistant to some or all of the standard antibiotics. This has created a demand for both new anti-microbial agents and diagnostic tests for this organism

Recently several novel approaches have been described which purport to follow global gene expression during infection (Chuang, S. et al., (1993); Mahan, M. J. et al., Science 259:686–688 (1993); Hensel, M. et al., Science 269:400–403 (1995). These new techniques have so far been demonstrated with gram negative pathogen infections, but not with infections with gram positives presumably because the much slower development of global transposon mutagenesis and suitable vectors needed for these strategies in these organisms, and in the case of that process described by Chuang, S. et al., J. Bacteriol. 175:2026–2036 (1993), the difficulty of isolating suitable quantities of bacterial RNA free of mammalian RNA derived from the infected tissue to furnish bacterial RNA labelled to sufficiently high specific activity.

The present invention employs a novel technology to determine gene expression in the pathogen at different stages of infection of the mammalian host. A novel aspect of this invention is the use of a suitably labelled oligonucleotide probe which anneals specifically to the bacterial ribosomal RNA in Northern blots of bacterial RNA preparations from infected tissue. Using the more abundant ribosomal RNA as a hybridization target greatly facilitates the optimization of a protocol to purify bacterial RNA of a suitable size and quantity for RT-PCR from infected tissue.

A suitable oligonucleotide useful for applying this method to genes expressed in *Staphylococcus aureus* is 5'-GCTCCTAAAAGGTTACTCCACCGGC-3' [SEQ ID NO:5].

Use of the technology of the present invention enables identification of bacterial genes transcribed during infection, inhibitors of which would have utility in anti-bacterial therapy. Specific inhibitors of such gene transcription or of the subsequent translation of the resultant mRNA or of the function of the corresponding expressed proteins would have utility in anti-bacterial therapy. This invention provides *Staphylococcus aureus* WCUH29 polynucleotides which are transcribed in infected tissue.

Clearly, there is a need for factors, such as the novel compounds of the invention, that have a present benefit of being useful to screen compounds for antibiotic activity. Such factors are also useful to determine their role in pathogenesis of infection, dysfunction and disease. There is also a need for identification and characterization of such factors and their antagonists and agonists which can play a role in preventing, ameliorating or correcting infections, dysfunctions or diseases.

The polypeptides of the invention have amino acid sequence homology to a known *Lactococcus lactis* ClpL protein (Genbank Accession Number: X62333).

SUMMARY OF THE INVENTION

It is an object of the invention to provide polypeptides that have been identified as novel clpL polypeptides by homology between the amino acid sequence set out in Table 1 [SEQ ID NO: 2] and a known amino acid sequence or sequences of other proteins such as Lactococcus lactis ClpL protein.

It is a further object of the invention to provide polynucleotides that encode clpL polypeptides, particularly polynucleotides that encode the polypeptide herein designated clpL.

In a particularly preferred embodiment of the invention, the polynucleotide comprises a region encoding clpL polypeptides comprising the sequence set out in Table 1 [SEQ ID NO: 1] which includes a fill length gene, or a variant thereof.

In another particularly preferred embodiment of the invention, there is a novel clpL protein from *Staphylococcus aureus* comprising the amino acid sequence of Table 1 [SEQ ID NO:2], or a variant thereof.

In accordance with another aspect of the invention, there is provided an isolated nucleic acid molecule encoding a mature polypeptide expressible by the *Staphylococcus aureus* WCUH 29 strain contained in the deposited stain.

As a further aspect of the invention, there are provided isolated nucleic acid molecules encoding clpL, particularly *Staphylococcus aureus* clpL, including mRNAs, cDNAs, genomic DNAs. Further embodiments of the invention include biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof and compositions comprising the same.

In accordance with another aspect of the invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization. Among the particularly preferred embodiments of the invention are naturally occurring allelic variants of clpL and polypeptides encoded thereby.

As another aspect of the invention, there are provided novel polypeptides of *Staphylococcus aureus* referred to herein as clpL as well as biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

Among the particularly preferred embodiments of the invention are variants of clpL polypeptide encoded by naturally occurring alleles of the clpL gene.

In a preferred emnbodiment of the invention, there are provided methods for producing the aforementioned clpL polypeptides.

In accordance with yet another aspect of the invention, there are provided inhibitors to such polypeptides, useful as antibacterial agents, including, for example, antibodies.

In accordance with certain preferred embodiments of the invention, there are provided products, compositions and methods for assessing clpL expression, treating disease, for example, disease, such as, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, darcryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis), assaying genetic variation, and administering a clpL polypeptide or polynucleotide to an organism to raise an immunological response against a bacteria, especially a *Staphylococcus aureus* bacteria.

In accordance with certain preferred embodiments of this and other aspects of the invention, there are provided polynucleotides that hybridize to clpL polynucleotide sequences, particularly under stringent conditions.

In certain preferred embodiments of the invention, there are provided antibodies against clpL polypeptides.

In other embodiments of the invention, there are provided methods for identifying compounds which bind to or otherwise interact with and inhibit or activate an activity of a polypeptide or polynucleotide of the invention comprising: contacting a polypeptide or polynucleotide of the invention with a compound to be screened under conditions to permit binding to or other interaction between the compound and the polypeptide or polynucleotide to assess the binding to or other interaction with the compound, such binding or interaction being associated with a second component capable of providing a detectable signal in response to the binding or interaction of the polypeptide or polynucleotide with the compound, and determining whether the compound binds to or otherwise interacts with and activates or inhibits an activity of the polypeptide or polynucleotide by detecting the presence or absence of a signal generated from the binding or interaction of the compound with the polypeptide or polynucleotide.

In accordance with yet another aspect of the invention, there are provided clpL agonists and antagonists, preferably bacteriostatic or bacteriocidal agonists and antagonists.

In a further aspect of the invention, there are provided compositions comprising a clpL polynucleotide or a clpL polypeptide for administration to a cell or to a multicellular organism.

Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

GLOSSARY

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"Host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous polynucleotide sequence.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A. M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D. W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM *J. Applied Math.*, 48: 1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S. F. et al., *J. Molec. Biol* 215: 403–410 (1990). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, Md. 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403–410 (1990). As an illustration, by a polynucleotide having a nucleotide sequence having at least, for example, 95% "identity" to a reference nucleotide sequence of SEQ ID NO: 1 it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence of SEQ ID NO: 1. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5 or 3 terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. Analogously, by a polypeptide having an amino acid sequence having at least, for example, 95% identity to a reference amino acid sequence of SEQ ID NO:2 is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids

of the reference amino acid of SEQ ID NO: 2. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

"Isolated" means altered "by the hand of man" from its natural state, i.e., if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotide(s)" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded regions, or a mixture of single- and double-stranded regions. In addition, "polynucleotide" as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term "polynucleotide(s)" also includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotide(s)" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term "polynucleotide(s)" as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including, for example, simple and complex cells. "Polynucleotide(s)" also embraces short polynucleotides often referred to as oligonucleotide(s).

"Polypeptide(s)" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds. "Polypeptide(s)" refers to both short chains, commonly referred to as peptides, oligopeptides and oligomers and to longer chains generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene encoded amino acids. "Polypeptide(s)" include those modified either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains, and the amino or carboxyl termini. Modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins, such as arginylation, and ubiquitination. See, for instance, *PROTEINS—STRUCTURE AND MOLECULAR PROPERTIES*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993) and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1–12 in *POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS*, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., *Meth. Enzymol.* 182:626–646 (1990) and Rattan et al., *Protein Synthesis: Posttranslational Modifications and Aging*, Ann. N.Y. Acad. Sci. 663: 48–62 (1992). Polypeptides may be branched or cyclic, with or without branching. Cyclic, branched and branched circular polypeptides may result from post-translational natural processes and may be made by entirely synthetic methods, as well.

"Variant(s)" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans.

DESCRIPTION OF THE INVENTION

The invention relates to novel clpL polypeptides and polynucleotides as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of a novel clpL of *Staphylococcus aureus*, which is related by amino acid sequence homology to Lactococcus lactis ClpL polypeptide (Lactococcus lactis clpL Genbank Accession Number: X62333). The invention relates especially to clpL having the nucleotide and amino acid sequences set out in Table 1 [SEQ ID NO: 1] and Table 1 [SEQ ID NO: 2] respectively, and to the clpL nucleotide sequences of the DNA in the deposited strain and amino acid sequences encoded thereby.

TABLE 1

clpL Polynucleotide and Polypeptide Sequences

(A) Sequences from *Staphylococcus aureus* clpL polynucleotide sequence [SEQ ID NO:1]

```
5'-1      ATGAATAACG GTTTTTTCAA TAGCGACTTT GATTCAATTT TTCGAAGAAT
  51      GATGCAAGAT ATGCAAGGTT CAAATCAAGT CGGTAACAAA AAGTACTATA
 101      TTAATGGTAA AGAAGTTTCA CCTGAAGAAC TAGCGCAACT CACACAACAA
 151      GGTAGCAATC AATCTGCTGA ACAAAGTGCG CAAGCTTTTC AACAAGCAGC
 201      ACAAAGACAA CAAGGGCAAC AAGGTGGCAA CGGCAATTAT TTAGAACAAA
 251      TTGGTCGTAA CCTTACGCAA GAAGCACGTG ACGGTTTATT AGATCCAGTC
 301      ATTGGTCGTG ATAAAGAAAT TCAAGAAACT GCTGAAGTCT TAAGTAGACG
 351      AACTAAAAAC AATCCTATAT TAGTTGGAGA AGCTGGTGTT GGTAAAACTG
 401      CGATTGTTGA AGGTTTAGCA CAGGCAATCG TTGAAGGAAA TGTACCAGCA
 451      GCAATCAAAG ACAAAGAAAT TATTTCTGTA GATATTTCAT CATTAGAAGC
 501      TGGAACGCAA TATCGTGGTG CTTTTGAAGA AAATATTCAA AAATTAATCG
 551      AAGGTGTTAA ATCTTCACAA AATGCCGTAC TATTCTTTGA TGAAATCCAT
 601      CAAATTATCG GTTCAGGTGC CACAGGAAGT GATTCAGGTA GCAAAGGGTT
 651      ATCTGATATT TTGAAACCTG CATTAAGTCG TGGTGAGATT TCTATTATTG
 701      GTGCAACAAC ACAAGATGAA TATCGAAACA ATATTCTTAA AGATGCTGCA
 751      TTAACGCGCA GATTTAATGA AGTGCTTGTT AATGAACCAA GCGCTAAAGA
 801      TACTGTTGAA ATTTTAAAAG GTATTCGCGA AAAATTCGAA GAACACCATC
 851      AAGTAAAATT ACCAGATGAC GTATTAAAAG CATGTGTTGA CTTATCAATT
 901      CAATATATTC CACAACGATT ATTACCAGAT AAAGCAATCG ATGTGTTAGA
 951      TATTACAGCA GCACATTTAT CTGCGCAAAG TCCAGCTGTC GATAAAGTTG
1001      AAACTGAAAA ACGAATTTCT GAATTAGAAA ATGATAAACG TAAAGCAGTA
1051      AGTGCTGAAG AATATAAAAA AGCTGACGAC ATTCAAAATG AAATCAAATC
1101      ATTACAAGAT AAATTAGAAA ACAGTAATGG TGAACATACT GCTGTTGCTA
1151      CAGTTCATGA TATTTCAGAT ACTATTCAAC GATTAACTGG CATTCCAGTT
1201      TCTCAAATGG ATGATAACGA TATTGAACGT TTAAAAAATA TTTCTAATCG
1251      TTTAAGAAGT AAAATCATAG GTCAAGATCA AGCTGTAGAA ATGGTTTCAC
1301      GTGCAATTCG CCGTAATCGC GCTGGGTTTG ATGACGGCAA CCGTCCAATT
1351      GGTAGTTTCT TATTTGTCGG CCCTACTGGT GTTGGTAAAA CAGAGCTTGC
1401      TAAACAATTA GCAATTGATC TATTTGGTAA TAAAGATGCA CTTATTCGAC
1451      TTGATATGAG TGAATATAGT GACACAACAG CTGTTTCAAA AATGATTGGT
1501      ACAACTGCTG GTTATGTCGG TTATGATGAC AATTCAAATA CGTTAACTGA
1551      AAAAGTACGC CGTAATCCAT ACTCAGTCAT TCTATTTGAT GAAATCGAAA
1601      AAGCAAATCC ACAAATTTTA ACATTGTTAT TACAAGTAAT GGATGATGGT
1651      AATTTGACTG ATGGTCAAGG TAATGTCATC AACTTTAAAA ATACAATTAT
1701      TATTTGTACA TCAAATGCTG GCTTTGGCAA TGGCAATGAC GCTGAAGAAA
1751      AAGATATTAT GCACGAAATG AAAAAATTCT TCCGCCCTGA ATTCCTTAAC
1801      CGCTTCAACG GCATCGTTGA ATTCTTACAT TTAGATAAAG ATGCATTGCA
1851      AGATATCGTC AACTTATTAT TAGACGATGT ACAAGTTACA TTAGATAAAA
1901      AAGGTATTAC GATGGACGTT TCTCAAGATG CGAAAGATTG GTTAATTGAA
1951      GAAGGCTATG ATGAAGAATT AGGTGCACGT CCATTAAGAC GTATTGTTGA
2001      ACAGCAAGTA CGTGACAAAA TTACAGATTA TTATCTAGAT CATACAGACG
2051      TTAAACATGT GGATATAGAT GTTGAGGATA ACGAATTAGT CGTAAAAGGT
2101      AAATAA-3'
```

(B) clpL polypeptide sequence deduced from the polynucleotide sequence in this table [SEQ ID NO:2].

```
NH2-1     MNNGFFNSDF DSIFRRMMQD MQGSNQVGNK KYYINGKEVS PEELAQLTQQ
  51      GSNQSAEQSA QAFQQAAQRQ QGQQGGNGNY LEQIGRNLTQ EARDGLLDPV
 101      IGRDKEIQET AEVLSRRTKN NPILVGEAGV GKTAIVEGLA QAIVEGNVPA
 151      AIKDKEIISV DISSLEAGTQ YRGAFEENIQ KLIEGVKSSQ NAVLFFDEIH
 201      QIIGSGATGS DSGSKGLSDI LKPALSRGEI SIIGATTQDE YRNNILKDAA
 251      LTRRFNEVLV NEPSAKDTVE ILKGIREKFE EHHQVKLPDD VLKACVDLSI
 301      QYIPQRLLPD KAIDVLDITA AHLSAQSPAV DKVETEKRIS ELENDKRKAV
 351      SAEEYKKADD IQNEIKSLQD KLENSNGEHT AVATVHDISD TIQRLTGIPV
 401      SQMDDNDIER LKNISNRLRS KIIGQDQAVE MVSRAIRRNR AGFDDGNRPI
 451      GSFLFVGPTG VGKTELAKQL AIDLFGNKDA LIRLDMSEYS DTTAVSKMIG
 501      TTAGYVGYDD NSNTLTEKVR RNPYSVILFD EIEKANPQIL TLLLQVMDDG
 551      NLTDGQGNVI NFKNTIIICT SNAGFGNGND AEEKDIMHEM KKFFRPEFLN
 601      RFNGIVEFLH LDKDALQDIV NLLLDDVQVT LDKKGITMDV SQDAKDWLIE
 651      EGYDEELGAR PLRRIVEQQV RDKITDYYLD HTDVKHVDID VEDNELVVKG
 701      K-COOH
```

(C) Polynucleotide sequence embodiments [SEQ ID NO:1].

```
X-(R1)n-1 ATGAATAACG GTTTTTTCAA TAGCGACTTT GATTCAATTT TTCGAAGAAT
      51  GATGCAAGAT ATGCAAGGTT CAAATCAAGT CGGTAACAAA AAGTACTATA
     101  TTAATGGTAA AGAAGTTTCA CCTGAAGAAC TAGCGCAACT CACACAACAA
     151  GGTAGCAATC AATCTGCTGA ACAAAGTGCG CAAGCTTTTC AACAAGCAGC
     201  ACAAAGACAA CAAGGGCAAC AAGGTGGCAA CGGCAATTAT TTAGAACAAA
```

TABLE 1-continued

clpL Polynucleotide and Polypeptide Sequences

```
      251 TTGGTCGTAA CCTTACGCAA GAAGCACGTG ACGGTTTATT AGATCCAGTC
      301 ATTGGTCGTG ATAAAGAAAT TCAAGAAACT GCTGAAGTCT TAAGTAGACG
      351 AACTAAAAAC AATCCTATAT TAGTTGGAGA AGCTGGTGTT GGTAAAACTG
      401 CGATTGTTGA AGGTTTAGCA CAGGCAATCG TTGAAGGAAA TGTACCAGCA
      451 GCAATCAAAG ACAAAGAAAT TATTTCTGTA GATATTTCAT CATTAGAAGC
      501 TGGAACGCAA TATCGTGGTG CTTTTGAAGA AAATATTCAA AAATTAATCG
      551 AAGGTGTTAA ATCTTCACAA AATGCCGTAC TATTCTTTGA TGAAATCCAT
      601 CAAATTATCG GTTCAGGTGC CACAGGAAGT GATTCAGGTA GCAAAGGGTT
      651 ATCTGATATT TTGAAACCTG CATTAAGTCG TGGTGAGATT TCTATTATTG
      701 GTGCAACAAC ACAAGATGAA TATCGAAACA ATATTCTTAA AGATGCTGCA
      751 TTAACGCGCA GATTTAATGA AGTGCTTGTT AATGAACCAA GCGCTAAAGA
      801 TACTGTTGAA ATTTTAAAAG GTATTCGCGA AAAATTCGAA GAACACCATC
      851 AAGTAAAATT ACCAGATGAC GTATTAAAAG CATGTGTTGA CTTATCAATT
      901 CAATATATTC CACAACGATT ATTACCAGAT AAAGCAATCG ATGTGTTAGA
      951 TATTACAGCA GCACATTTAT CTGCGCAAAG TCCAGCTGTC GATAAAGTTG
     1001 AAACTGAAAA ACGAATTTCT GAATTAGAAA ATGATAAACG TAAAGCAGTA
     1051 AGTGCTGAAG AATATAAAAA AGCTGACGAC ATTCAAAATG AAATCAAATC
     1101 ATTACAAGAT AAATTAGAAA ACAGTAATGG TGAACATACT GCTGTTGCTA
     1151 CAGTTCATGA TATTTCAGAT ACTATTCAAC GATTAACTGG CATTCCAGTT
     1201 TCTCAAATGG ATGATAACGA TATTGAACGT TTAAAAAATA TTTCTAATCG
     1251 TTTAAGAAGT AAAATCATAG GTCAAGATCA AGCTGTAGAA ATGGTTTCAC
     1301 GTGCAATTCG CCGTAATCGC GCTGGGTTTG ATGACGGCAA CCGTCCAATT
     1351 GGTAGTTTCT TATTTGTCGG CCCTACTGGT GTTGGTAAAA CAGAGCTTGC
     1401 TAAACAATTA GCAATTGATC TATTTGGTAA TAAAGATGCA CTTATTCGAC
     1451 TTGATATGAG TGAATATAGT GACACAACAG CTGTTTCAAA AATGATTGGT
     1501 ACAACTGCTG GTTATGTCGG TTATGATGAC AATTCAAATA CGTTAACTGA
     1551 AAAAGTACGC CGTAATCCAT ACTCAGTCAT TCTATTTGAT GAAATCGAAA
     1601 AAGCAAATCC ACAAATTTTA ACATTGTTAT TACAAGTAAT GGATGATGGT
     1651 AATTTGACTG ATGGTCAAGG TAATGTCATC AACTTTAAAA ATACAATTAT
     1701 TATTTGTACA TCAAATGCTG GCTTTGGCAA TGGCAATGAC GCTGAAGAAA
     1751 AAGATATTAT GCACGAAATG AAAAAATTCT TCCGCCCTGA ATTCCTTAAC
     1801 CGCTTCAACG GCATCGTTGA ATTCTTACAT TTAGATAAAG ATGCATTGCA
     1851 AGATATCGTC AACTTATTAT TAGACGATGT ACAAGTTACA TTAGATAAAA
     1901 AAGGTATTAC GATGGACGTT TCTCAAGATG CGAAAGATTG GTTAATTGAA
     1951 GAAGGCTATG ATGAAGAATT AGGTGCACGT CCATTAAGAC GTATTGTTGA
     2001 ACAGCAAGTA CGTGACAAAA TTACAGATTA TTATCTAGAT CATACAGACG
     2051 TTAAACATGT GGATATAGAT GTTGAGGATA ACGAATTAGT CGTAAAAGGT
     2101 AAATAA-(R2)n-Y
```

(D) Polypeptide sequence embodiments [SEQ ID NO:2].

```
X-(R1)n-1  MNNGFFNSDF DSIFRRMMQD MQGSNQVGNK KYYINGKEVS PEELAQLTQQ
       51  GSNQSAEQSA QAFQQAAQRQ QGQQGGNGNY LEQIGRNLTQ EARDGLLDPV
      101  IGRDKEIQET AEVLSRRTKN NPILVGEAGV GKTAIVEGLA QAIVEGNVPA
      151  AIKDKEIISV DISSLEAGTQ YRGAFEENIQ KLIEGVKSSQ NAVLFFDEIH
      201  QIIGSGATGS DSGSKGLSDI LKPALSRGEI SIIGATTQDE YRNNILKDAA
      251  LTRRFNEVLV NEPSAKDTVE ILKGIREKFE EHHQVKLPDD VLKACVDLSI
      301  QYIPQRLLPD KAIDVLDITA AHLSAQSPAV DKVETEKRIS ELENDKRKAV
      351  SAEEYKKADD IQNEIKSLQD KLENSNGEHT AVATVHDISD TIQRLTGIPV
      401  SQMDDNDIER LKNISNRLRS KIIGQDQAVE MVSRAIRRNR AGFDDGNRPI
      451  GSFLFVGPTG VGKTELAKQL AIDLFGNKDA LIRLDMSEYS DTTAVSKMIG
      501  TTAGYVGYDD NSNTLTEKVR RNPYSVILFD EIEKANPQIL TLLLQVMDDG
      551  NLTDGQGNVI NFKNTIIICT SNAGFGNGND AEEKDIMHEM KKFFRPEFLN
      601  RFNGIVEFLH LDKDALQDIV NLLLDDVQVT LDKKGITMDV SQDAKDWLIE
      651  EGYDEELGAR PLRRIVEQQV RDKITDYYLD HTDVKHVDID VEDNELVVKG
      701  K-(R2)n-Y
```

Deposited Materials

A deposit containing a *Staphylococcus aureus* WCUH 29 strain has been deposited with the National Collections of Industrial and Marine Bacteria Ltd. (herein "NCIMB"), 23 St. Machar Drive, Aberdeen AB2 1RY, Scotland on Sep. 11, 1995 and assigned NCIMB Deposit No. 40771, and referred to as *Staphylococcus aureus* WCUH29 on deposit. The *Staphylococcus aureus* strain deposit is referred to herein as "the deposited strain" or as "the DNA of the deposited strain."

The deposited strain contains the full length clpL gene. The sequence of the polynucleotides contained in the deposited strain, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

The deposit of the deposited strain has been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposited strain is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that required under 35 U.S.C. §112.

A license may be required to make, use or sell the deposited strain, and compounds derived therefrom, and no such license is hereby granted.

Polypeptides

The polypeptides of the invention include the polypeptide of Table 1 [SEQ ID NO:2] (in particular the mature polypeptide) as well as polypeptides and fragments, particularly those which have the biological activity of clpL, and also those which have at least 70% identity to the polypeptide of Table 1 [SEQ ID NO:2] or the relevant portion, preferably at least 80% identity to the polypeptide of Table 1 [SEQ ID NO:2], and more preferably at least 90% simi-

larity (more preferably at least 90% identity) to the polypeptide of Table 1 [SEQ ID NO:2] and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the polypeptide of Table 1 [SEQ ID NO:2] and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids.

The invention also includes polypeptides of the formula set forth in Table 1 (D) wherein, at the amino terminus, X is hydrogen, and at the carboxyl terminus, Y is hydrogen or a metal, R_1 and R_2 is any amino acid residue, and n is an integer between 1 and 1000. Any stretch of amino acid residues denoted by either R group, where R is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer.

A fragment is a variant polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned polypeptides. As with clpL polypeptides fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region, a single larger polypeptide.

Preferred fragments include, for example, truncation polypeptides having a portion of the amino acid sequence of Table 1 [SEQ ID NO:2], or of variants thereof, such as a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus. Degradation forms of the polypeptides of the invention in a host cell, particularly a *Staphylococcus aureus*, are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

Also preferred are biologically active fragments which are those fragments that mediate activities of clpL, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those fragments that are antigenic or immunogenic in an animal, especially in a human. Particularly preferred are fragments comprising receptors or domains of enzymes that confer a function essential for viability of *Staphylococcus aureus* or the ability to initiate, or maintain cause disease in an individual, particularly a human.

Variants that are fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides of the invention.

Polynucleotides

Another aspect of the invention relates to isolated polynucleotides, including the fill length gene, that encode the clpL polypeptide having the deduced amino acid sequence of Table 1 [SEQ ID NO:2] and polynucleotides closely related thereto and variants thereof

Using the information provided herein, such as the polynucleotide sequence set out in Table 1 [SEQ ID NO: 1], a polynucleotide of the invention encoding clpL polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and sequencing chromosomal DNA fragments from bacteria using *Staphylococcus aureus* WCUH 29 cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide sequence of the invention, such as the sequence given in Table 1 [SEQ ID NO:1], typically a library of clones of chromosomal DNA of *Staphylococcus aureus* WCUH 29 in *E. coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent conditions. By sequencing the individual clones thus identified with sequencing primers designed from the original sequence it is then possible to extend the sequence in both directions to determine the full gene sequence. Conveniently, such sequencing is performed using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E. F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). (see in particular Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Illustrative of the invention, the polynucleotide set out in Table 1 [SEQ ID NO:1] was discovered in a DNA library derived from *Staphylococcus aureus* WCUH 29.

The DNA sequence set out in Table 1 [SEQ ID NO: 1] contains an open reading frame encoding a protein having about the number of amino acid residues set forth in Table 1 [SEQ ID NO:2] with a deduced molecular weight that can be calculated using amino acid residue molecular weight values well known in the art. The polynucleotide of SEQ ID NO: 1, between nucleotide number 1 through number 2103 encodes the polypeptide of SEQ ID NO:2. The stop codon begins at nucleotide number 2104 of SEQ ID NO:1.

The clpL polypeptide of the invention is structurally related to other proteins of the ATP dependent proteases family, as shown by the results of sequencing the DNA encoding clpL of the deposited strain. The protein exhibits greatest homology to *Laetococcus lactis* ClpL protein among known proteins. The clpL polypeptide of Table 1 [SEQ ID NO:2] has about 55% identity over its entire length and about 73% similarity over its entire length with the amino acid sequence of *Lactococcus lactis* ClpL polypeptide.

The invention provides a polynucleotide sequence identical over its entire length to the coding sequence in Table 1 [SEQ ID NO: 1]. Also provided by the invention is the coding sequence for the mature polypeptide or a fragment thereof, by itself as well as the coding sequence for the mature polypeptide or a fragment in reading frame with other coding sequence, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence. The polynucleotide may also contain non coding sequences, including for example, but not limited to non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences, termination signals, ribosome binding sites, sequences that stabilize mRNA, introns, polyadenylation signals, and additional coding sequence which encode additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz et al., *Proc. Natl. Acad. Sci.*, USA 86: 821–824 (1989), or an HA tag (Wilson et al., *Cell* 37: 767 (1984). Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

A preferred embodiment of the invention is the polynucleotide of comprising nucleotide 1 to 2103 set forth in SEQ ID NO:1 of Table 1 which encodes the clpL polypeptide.

The invention also includes polynucleotides of the formula set forth in Table 1 (C) wherein, at the 5' end of the molecule, X is hydrogen, and at the 3' end of the molecule, Y is hydrogen or a metal, R_1 and R_2 is any nucleic acid residue, and n is an integer between 1 and 1000. Any stretch of nucleic acid residues denoted by either R group, where R is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer.

The term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides that include a sequence encoding a polypeptide of the invention, particularly a bacterial polypeptide and more particularly a polypeptide of the *Staphylococcus aureus* clpL having the amino acid sequence set out in Table 1 [SEQ ID NO:2]. The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by integrated phage or an insertion sequence or editing) together with additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the polynucleotides described herein that encode for variants of the polypeptide having the deduced amino acid sequence of Table 1 [SEQ ID NO:2]. Variants that are fragment of the polynucleotides of the invention may be used to synthesize full-length polynucleotides of the invention.

Further particularly preferred embodiments are polynucleotides encoding clpL variants, that have the amino acid sequence of clpL polypeptide of Table 1 [SEQ ID NO:2] in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, that do not alter the properties and activities of clpL.

Further preferred embodiments of the invention are polynucleotides that are at least 70% identical over their entire length to a polynucleotide encoding clpL polypeptide having the amino acid sequence set out in Table 1 [SEQ ID NO:2], and polynucleotides that are complementary to such polynucleotides. Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 80% identical over its entire length to a polynucleotide encoding clpL polypeptide of the deposited strain and polynucleotides complementary thereto. In this regard, polynucleotides at least 90% identical over their entire length to the same are particularly preferred, and among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

Preferred embodiments are polynucleotides that encode polypeptides that retain substantially the same biological function or activity as the mature polypeptide encoded by the DNA of Table 1 [SEQ ID NO: 1].

The invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. An example of stringent hybridization conditions is overnight incubation at 42° C. in a solution comprising: 50% formamide, 5×SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5×Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1×SSC at about 65° C. Hybridization and wash conditions are well known and exemplified in Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein.

The invention also provides a polynucleotide consisting essentially of a polynucleotide sequence obtainable by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in SEQ ID NO:1 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence set forth in SEQ ID NO:1 or a fragment thereof; and isolating said DNA sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers described elsewhere herein.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, polynucleotides of the invention as discussed above, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding clpL and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the clpL gene.

Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may have at least 50 bases. Particularly preferred probes will have at least 30 bases and will have 50 bases or less.

For example, the coding region of the clpL gene may be isolated by screening using the DNA sequence provided in SEQ ID NO: 1 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for disease, particularly human disease, as further discussed herein relating to polynucleotide assays.

Polynucleotides of the invention that are oligonucleotides derived from the sequences of SEQ ID NOS:1 and/or 2 may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The invention also provides polynucleotides that may encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case in vivo, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the invention may encode a mature protein a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

Vectors, Host Cells, Expression

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis et al, *BASIC METHODS MOLECULAR BIOLOGY*, (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, enterococci *E. coli*, streptomyces and *Bacillus subtilis* cells; fungal cells, such as yeast cells and Aspergillus cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, e.g, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, (supra).

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Diagnostic Assays

This invention is also related to the use of the clpL polynucleotides of the invention for use as diagnostic reagents. Detection of clpL in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of a disease. Eukaryotes (herein also "individual(s)"), particularly mammals, and especially humans, particularly those infected or suspected to be infected with an organism comprising the clpL gene may be detected at the nucleic acid level by a variety of techniques.

Nucleic acids for diagnosis may be obtained from an infected individual's cells and tissues, such as bone, blood, muscle, cartilage, and skin. Genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification technique prior to analysis. RNA or cDNA may also be used in the same ways. Using amplification, characterization of the species and strain of prokaryote present in an individual, may be made by an analysis of the genotype of the prokaryote gene. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the genotype of a reference sequence. Point mutations can be identified by hybridizing amplified DNA to labeled clpL polynucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in the electrophoretic mobility of the DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g, Myers et al., *Science*, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or a chemical cleavage method. See, e.g, Cotton et al., *Proc. Natl. Acad. Sci.*, USA, 85: 4397–4401 (1985).

Cells carrying mutations or polymorphisms in the gene of the invention may also be detected at the DNA level by a variety of techniques, to allow for serotyping, for example. For example, RT-PCR can be used to detect mutations. It is particularly preferred to used RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA or cDNA may also be used for the same purpose, PCR or RT-PCR. As an example, PCR primers complementary to a nucleic acid encoding clpL can be used to identify and analyze mutations. Examples of representative primers are shown below in Table 2.

TABLE 2

Primers for amplification of clpL polynucleotides

SEQ ID NO	PRIMER SEQUENCE	
3	5'-ATGAATAACG GTTTTTTCAA TAGC-3'	[SEQ ID NO:3]
4	5'-TTATTTACCT TTTACGACTA ATTC-3'	[SEQ ID NO:4]

The invention further provides these primers with 1, 2, 3 or 4 nucleotides removed from the 5' and/or the 3' end. These primers may be used for, among other things, amplifying clpL DNA isolated from a sample derived from an individual. The primers may be used to amplify the gene isolated from an infected individual such that the gene may then be

subject to various techniques for elucidation of the DNA sequence. In this way, mutations in the DNA sequence may be detected and used to diagnose infection and to serotype and/or classify the infectious agent.

The invention further provides a process for diagnosing, disease, preferably bacterial infections, more preferably infections by *Staphylococcus aureus*, and most preferably disease, such as, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, darcryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis), comprising determining from a sample derived from an individual a increased level of expression of polynucleotide having the sequence of Table 1 [SEQ ID NO: 1]. Increased or decreased expression of clpL polynucleotide can be measured using any on of the methods well known in the art for the quantation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods.

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of clpL protein compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a clpL protein, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Antibodies

The polypeptides of the invention or variants thereof, or cells expressing them can be used as an immunogen to produce antibodies immunospecific for such polypeptides. "Antibodies" as used herein includes monoclonal and polyclonal antibodies, chimeric, single chain, simianized antibodies and humanized antibodies, as well as Fab fragments, including the products of an Fab immunolglobulin expression library.

Antibodies generated against the polypeptides of the invention can be obtained by administering the polypeptides or epitope-bearing fragments, analogues or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature* 256: 495–497 (1975); Kozbor et al., *Immunology Today* 4: 72 (1983); Cole et al., pg. 77–96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Pat. No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies.

Alternatively phage display technology may be utilized to select antibody genes with binding activities towards the polypeptide either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-clpL or from naive libraries (McCafferty, J. et al., (1990), *Nature* 348, 552–554; Marks, J. et al., (1992) *Biotechnology* 10, 779–783). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. et al., (1991) *Nature* 352, 624–628).

If two antigen binding domains are present each domain may be directed against a different epitope—termed 'bispecific' antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides to purify the polypeptides by affinity chromatography.

Thus, among others, antibodies against clpL- polypeptide may be employed to treat infections, particularly bacterial infections and especially disease, such as, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, darcryocystitis), kidney and urinary tract (e.g., epididymitis, intarenal and perinephric abscess, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis).

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants that form a particular aspect of this invention. The term "antigenically equivalent derivative" as used herein encompasses a polypeptide or its equivalent which will be specifically recognized by certain antibodies which, when raised to the protein or polypeptide according to the invention, interfere with the immediate physical interaction between pathogen and mammalian host. The term "immunologically equivalent derivative" as used herein encompasses a peptide or its equivalent which when used in a suitable formulation to raise antibodies in a vertebrate, the antibodies act to interfere with the immediate physical interaction between pathogen and mammalian host.

The polypeptide, such as an antigenically or immunologically equivalent derivative or a fusion protein thereof is used as an antigen to immunize a mouse or other animal such as a rat or chicken. The fusion protein may provide stability to the polypeptide. The antigen may be associated, for example by conjugation, with an immunogenic carrier protein for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively a multiple antigenic peptide comprising multiple copies of the protein or polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized"; where the complimentarily determining region(s) of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones, P. et al. (1986), *Nature* 321, 522–525 or Tempest et al.,(1991) *Biotechnology* 9, 266–273.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff et al., *Hum Mol Genet* 1992, 1:363, Manthorpe et al., *Hum. Gene Ther.* 1963:4, 419), delivery of DNA complexed with specific protein carriers (Wu et al., *J Biol Chem.* 1989: 264,16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, PNAS USA, 1986:83,9551), encapsulation of DNA in various forms of

liposomes (Kaneda et al., *Science* 1989:243,375), particle bombardment (Tang et at, *Nature* 1992, 356:152, Eisenbraun et al., *DNA Cell Biol* 1993, 12:791) and in vivo infection using cloned retroviral vectors (Seeger et al., *PNAS USA* 1984:81,5849).

Antagonists and Agonists—Assays and Molecules

Polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, e.g, Coligan et al., *Current Protocols in Immunology* 1(2): Chapter 5 (1991).

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of clpL polypeptides or polynucleotides, particularly those compounds that are bacteriostatic and/or bacteriocidal. The method of screening may involve high-throughput techniques. For example, to screen for agonists or antagoists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising clpL polypeptide and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be a clpL agonist or antagonist. The ability of the candidate molecule to agonize or antagonize the clpL polypeptide is reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Molecules that bind gratuitously, i.e., without inducing the effects of clpL polypeptide are most likely to be good antagonists. Molecules that bind well and increase the rate of product production from substrate are agonists. Detection of the rate or level of production of product from substrate may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric labeled substrate converted into product, a reporter gene that is responsive to changes in clpL polynucleotide or polypeptide activity, and binding assays known in the art.

Another example of an assay for clpL antagonists is a competitive assay that combines clpL and a potential antagonist with clpL-binding molecules, recombinant clpL binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. The clpL molecule can be labeled, such as by radioactivity or a colorimetric compound, such that the number of clpL molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide or polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing clpL-induced activities, thereby preventing the action of clpL by excluding clpL from binding.

Potential antagonists include a small molecule that binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem* 56: 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*, CRC Press, Boca Raton, Fla. (1988), for a description of these molecules). Preferred potential antagonists include compounds related to and variants of clpL.

Each of the DNA sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide or inhibitor of the invention to interfere with the initial physical interaction between a pathogen and mammalian host responsible for sequelae of infection. In particular the molecules of the invention may be used: in the prevention of adhesion of bacteria, in particular gram positive bacteria, to mammalian extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; to block clpL protein-mediated mammalian cell invasion by, for example, initiating phosphorylation of mammalian tyrosine kinases (Rosenshine et al, *Infect. Immun.* 60:2211 (1992); to block bacterial adhesion between mammalian extracellular matrix proteins and bacterial clpL proteins that mediate tissue damage and; to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or. by other surgical techniques.

The antagonists and agonists of the invention may be employed, for instance, to inhibit and treat disease, such as, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic absces, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, darcryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric absces, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis).

Helicobacter pylori (herein *H. pylori*) bacteria infect the stomachs of over one-third of the world's population causing stomach cancer, ulcers, and gastritis (International Agency for Research on Cancer (1994) Schistosomes, Liver Flukes and *Helicobacter Pylori* (International Agency for Research on Cancer, Lyon, France; http://www.uicc.ch/ecp/ecp2904.htm). Moreover, the international Agency for Research on Cancer recently recognized a cause-and-effect relationship between *H. pylori* and gastric adenocarcinoma, classifying the bacterium as a Group I (definite) carcinogen. Preferred antimicrobial compounds of the invention (agonists and antagonists of clpL) found using screens provided by the invention, particularly broad-spectrum antibiotics, should be useful in the treatment of *H. pylori* infection. Such treatment should decrease the advent of *H. pylori*-induced cancers, such as gastrointestinal carcinoma. Such treatment should also cure gastric ulcers and gastritis.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal which comprises inoculating the individual with clpL, or a fragment or variant thereof, adequate to produce antibody and/ or T cell immune response to

protect said individual from infection, particularly bacterial infection and most particularly *Staphylococcus aureus* infection. Also provided are methods whereby such immunological response slows bacterial replication. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector to direct expression of clpL, or a fragment or a variant thereof, for expressing clpL, or a fragment or a variant thereof in vivo in order to induce an immunological response, such as, to produce antibody and/ or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual from disease, whether that disease is already established within the individual or not. One way of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise.

Such nucleic acid vector may comprise DNA, RNA, a modified nucleic acid, or a DNA/RNA hybrid.

A further aspect of the invention relates to an immunological composition which, when introduced into an individual capable or having induced within it an immunological response, induces an immunological response in such individual to a clpL or protein coded therefrom, wherein the composition comprises a recombinant clpL or protein coded therefrom comprising DNA which codes for and expresses an antigen of said clpL or protein coded therefrom. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity or cellular immunity such as that arising from CTL or CD4+T cells.

A clpL polypeptide or a fragment thereof may be fused with co-protein which may not by itself produce antibodies, but is capable of stabilizing the first protein and producing a fused protein which will have immunogenic and protective properties. Thus fused recombinant protein, preferably further comprises an antigenic co-protein, such as lipoprotein D from *Hemophilus influenzae*, Glutathione-S-transferase (GST) or beta-galactosidase, relatively large co-proteins which solubilize the protein and facilitate production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system. The co-protein may be attached to either the amino or carboxy terminus of the first protein.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. et al. *Science* 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof which have been shown to encode non-variable regions of bacterial cell surface proteins in DNA constructs used in such genetic immunization experiments in animal models of infection with *Staphylococcus aureus* will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value from the requisite organ of the animal successfully resisting or clearing infection for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *Staphylococcus aureus* infection, in mammals, particularly humans.

The polypeptide may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria, for example by blocking adherence of bacteria to damaged tissue. Examples of tissue damage include wounds in skin or connective tissue caused, e.g., by mechanical, chemical or thermal damage or by implantation of indwelling devices, or wounds in the mucous membranes, such as the mouth, mammary glands, urethra or vagina.

The invention also includes a vaccine formulation which comprises an immunogenic recombinant protein of the invention together with a suitable carrier. Since the protein may be broken down in the stomach, it is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation insotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

While the invention has been described with reference to certain clpL protein, it is to be understood that this covers fragments of the naturally occurring protein and similar proteins with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant protein.

Compositions, Kits and Administration

The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or their agonists or antagonists. The polypeptides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops,

mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation.

For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

In-dwelling devices include surgical implants, prosthetic devices and catheters, i.e., devices that are introduced to the body of an individual and remain in position for an extended time. Such devices include, for example, artificial joints, heart valves, pacemakers, vascular grafts, vascular catheters, cerebrospinal fluid shunts, urinary catheters, continuous ambulatory peritoneal dialysis (CAPD) catheters.

The composition of the invention may be administered by injection to achieve a systemic effect against relevant bacteria shortly before insertion of an in-dwelling device. Treatment may be continued after surgery during the in-body time of the device. In addition, the composition could also be used to broaden perioperative cover for any surgical technique to prevent bacterial wound infections, especially *Staphylococcus aureus* wound infections.

Many orthopaedic surgeons consider that humans with prosthetic joints should be considered for antibiotic prophylaxis before dental treatment that could produce a bacteremia. Late deep infection is a serious complication sometimes leading to loss of the prosthetic joint and is accompanied by significant morbidity and mortality. It may therefore be possible to extend the use of the active agent as a replacement for prophylactic antibiotics in this situation.

In addition to the therapy described above, the compositions of this invention may be used generally as a wound treatment agent to prevent adhesion of bacteria to matrix proteins exposed in wound tissue and for prophylactic use in dental treatment as an alternative to, or in conjunction with, antibiotic prophylaxis.

Alternatively, the composition of the invention may be used to bathe an indwelling device immediately before insertion. The active agent will preferably be present at a concentration of 1 μg/ml to 10 mg/ml for bathing of wounds or indwelling devices.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.5–5 microgram/kg of antigen, and such dose is preferably administered 1–3 times and with an interval of 1–3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

Each reference disclosed herein is incorporated by reference herein in its entirety. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety.

EXAMPLES

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples are illustrative, but do not limit the invention.

Example 1

Strain Selection, Library Production and Sequencing

The polynucleotide having the DNA sequence given in SEQ ID NO:1 was obtained from a library of clones of chromosomal DNA of *Staphylococcus aureus* in *E. coli*. The sequencing data from two or more clones containing overlapping *Staphylococcus aureus* DNAs was used to construct the contiguous DNA sequence in SEQ ID NO:1. Libraries may be prepared by routine methods, for example:
Methods 1 and 2 below.

Total cellular DNA is isolated from *Staphylococcus aureus* WCUH 29 according to standard procedures and size-fractionated by either of two methods.

Method 1

Total cellular DNA is mechanically sheared by passage through a needle in order to size-fractionate according to standard procedures. DNA fragments of up to 11 kbp in size are rendered blunt by treatment with exonuclease and DNA polymerase, and EcoRI linkers added. Fragments are ligated into the vector Lambda ZapII that has been cut with EcoRI, the library packaged by standard procedures and *E. coli* infected with the packaged library. The library is amplified by standard procedures.

Method 2

Total cellular DNA is partially hydrolyzed with a one or a combination of restriction enzymes appropriate to generate a series of fragments for cloning into library vectors (e.g., RsaI, PalI, AluI, Bshl235I), and such fragments are size-fractionated according to standard procedures. EcoRI linkers are ligated to the DNA and the fragments then ligated into the vector Lambda ZapII that have been cut with EcoRI, the library packaged by standard procedures, and *E. coli* infected with the packaged library. The library is amplified by standard procedures.

Example 2

The Determination of Expression During Infection of a Gene From *Staphylococcus aureus*

Necrotic fatty tissue from a four day groin infection of *Staphylococcus aureus* WCUH29 in the mouse is efficiently disrupted and processed in the presence of chaotropic agents and RNAase inhibitor to provide a mixture of animal and bacterial RNA. The optimal conditions for disruption and processing to give stable preparations and high yields of bacterial RNA are followed by the use of hybridisation to a radiolabelled oligonucleotide specific to *Staphylococcus aureus* 16S RNA on Northern blots. The RNAase free, DNAase free, DNA and protein free preparations of RNA obtained are suitable for Reverse Transcription PCR (RT-PCR) using unique primer pairs designed from the sequence of each gene of *Staphylococcus aureus* WCUH29.

a) Isolation of tissue infected with *Staphylococcus aureus* WCUH29 from a mouse animal model of infection

10 ml. volumes of sterile nutrient broth (No.2 Oxoid) are seeded with isolated, individual colonies of *Staphylococcus aureus* WCUH29 from an agar culture plate. The cultures are incubated aerobically (static culture) at 37° C. for 16–20

hours . 4 week old mice (female,18 g–22 g, strain MF1) are each infected by subcutaneous injection of 0.5 ml. of this broth culture of *Staphylococcus aureus* WCUH29 (diluted in broth to approximately 108 cfu/ml.) into the anterior, right lower quadrant (groin area). Mice should be monitored regularly during the first 24 hours after infection, then daily until termination of study. Animals with signs of systemic infection, i.e. lethargy, ruffled appearance, isolation from group, should be monitored closely and if signs progress to moribundancy, the animal should be culled immediately.

Visible external signs of lesion development will be seen 24–48h after infection. Examination of the abdomen of the animal will show the raised outline of the abscess beneath the skin. The localized lesion should remain in the right lower quadrant, but may occasionally spread to the left lower quadrant, and superiorly to the thorax. On occasions, the abscess may rupture through the overlying skin layers. In such cases the affected animal should be culled immediately and the tissues sampled if possible. Failure to cull the animal may result in the necrotic skin tissue overlying the abscess being sloughed off, exposing the abdominal muscle wall.

Approximately 96 hours after infection, animals are killed using carbon dioxide asphyxiation. To minimize delay between death and tissue processing/storage, mice should be killed individually rather than in groups. The dead animal is placed onto its back and the fur swabbed liberally with 70% alcohol. An initial incision using scissors is made through the skin of the abdominal left lower quadrant, travelling superiorly up to, then across the thorax. The incision is completed by cutting inferiorly to the abdominal lower right quadrant. Care should be taken not to penetrate the abdominal wall. Holding the skin flap with forceps, the skin is gently pulled way from the abdomen. The exposed abscess, which covers the peritoneal wall but generally does not penetrate the muscle sheet completely, is excised, taking care not to puncture the viscera.

The abscess/muscle sheet and other infected tissue may require cutting in sections, prior to flash-freezing in liquid nitrogen, thereby allowing easier storage in plastic collecting vials.

b) Isolation of *Staphylococcus aureus* WCUH29 RNA from infected tissue samples

4–6 infected tissue samples(each approx 0.5–0.7g) in 2 ml screw-cap tubes are removed from −8° C. storage into a dry ice ethanol bath In a microbiological safety cabinet the samples are disrupted individually whilst the remaining samples are kept cold in the dry ice ethanol bath. To disrupt the bacteria within the tissue sample 1 ml of TRIzol Reagent (Gibco BRL, Life Technologies) is added followed by enough 0.1 mm zirconia/silica beads to almost fill the tube, the lid is replaced taking care not to get any beads into the screw thread so as to ensure a good seal and eliminate aerosol generation. The sample is then homogenised in a Mini-BeadBeater Type BX-4 (Biospec Products). Necrotic fatty tissue isstrain treated for 100 seconds at 5000 rpm in order to achieve bacterial lysis. In vivo grown bacteria require longer treatment than in vitro grown *Staphylococcus aureus* which are disrupted by a 30 second bead-beat.

After bead-beating the tubes are chilled on ice before opening in a fume-hood as heat generated during disruption may degrade the TRIzol and release cyanide.

200 microlitres of chloroform is then added and the tubes shaken by hand for 15 seconds to ensure complete mixing. After 2–3 minutes at room temperature the tubes are spun down at 12,000×g, 4° C. for 15minutes and RNA extraction is then continued according to the method given by the manufacturers of TRIzol Reagent i.e.: The aqueous phase, approx 0.6 ml, is transferred to a sterile eppendorf tube and 0.5 ml of isopropanol is added. After 10 minutes at room temperature the samples are spun at 12,000×g, 4° C. for 10 minutes. The supernatant is removed and discarded then the RNA pellet is washed with 1 ml 75% ethanol. A brief vortex is used to mix the sample before centrifuging at 7,500×g, 4° C. for 5 minutes. The ethanol is removed and the RNA pellet dried under vacuum for no more than 5 minutes. Samples are then resuspended by repeated pipetting in 100 microlitres of DEPC treated water, followed by 5–10 minutes at 55° C. Finally, after at least 1 minute on ice, 200 units of Rnasin (Promega) is added.

RNA preparations are stored at −80° C. for up to one month. For longer term storage the RNA precipitate can be stored at the wash stage of the protocol in 75% ethanol for at least one year at −20° C.

Quality of the RNA isolated is assessed by running samples on 1% agarose gels. 1×TBE gels stained with ethidium bromide are used to visualise total RNA yields. To demonstrate the isolation of bacterial RNA from the infected tissue 1×MOPS, 2.2 M formaldehyde gels are run and vacuum blotted to Hybond-N (Amersham). The blot is then hybridised with a 32 P labelled oligonucletide probe specific to 16s rRNA of *Staphylococcus aureus* (K.Greisen, M. Loeffelholz, A. Purohit and D. Leong. J.Clin. (1994) Microbiol. 32 335–351). An oligonucleotide of the sequence: 5'-GCTCCTAAAAGGTTACTCCACCGGC-3'[SEQ ID NO:5] is used as a probe. The size of the hybridizing band is compared to that of control RNA isolated from in vitro grown *Staphylococcus aureus* WCUH29 in the Northern blot. Correct sized bacterial 16s rRNA bands can be detected in total RNA samples which show extensive degradation of the mammalian RNA when visualised on TBE gels.

c) The removal of DNA from *Staphylococcus aureus* WCUH29-derived RNA

DNA was removed from 73 microlitre samples of RNA by a 15 minute treatment on ice with 3 units of DNAaseI, amplification grade (Gibco BRL, Life Technologies) in the buffer supplied with the addition of 200 units of Rnasin (Promega) in a final volume of 90 microlitres.

The DNAase was inactivated and removed by treatment with TRIzol LS Reagent (Gibco BRL, Life Technologies) according to the manufacturers protocol. DNAase treated RNA was resuspended in 73 microlitres of DEPC treated water with the addition of Rnasin as described in Method 1.

d) The preparation of cDNA from RNA samples derived from infected tissue

10 microlitre samples of DNAase treated RNA are reverse transcribed using a SuperScript Preamplification System for First Strand cDNA Synthesis kit (Gibco BRL, Life Technologies) according to the manufacturers instructions. 1 nanogram of random hexamers is used to prime each reaction. Controls without the addition of SuperScriptII reverse transcriptase are also run. Both +/−RT samples are treated with RNaseH before proceeding to the PCR reaction.

e) The use of PCR to determine the presence of a bacterial cDNA species PCR reactions are set up on ice in 0.2 ml tubes by adding the following components: 45 microlitres PCR SUPERMIX (Gibco BRL, Life Technologies); 1 microlitre 50 mM MgCl2, to adjust final concentration to 2.5 mM; 1 microlitre PCR primers(optimally 18–25 basepairs in length and designed to possess similar annealing temperatures), each primer at 10 mM initial concentration; and 2 microlitres cDNA.

PCR reactions are run on a Perkin Elmer GeneAmp PCR System 9600 as follows: 5 minutes at 95° C., then 50 cycles

of 30 seconds each at 94° C., 42° C. and 72° C. followed by 3 minutes at 72° C. and then a hold temperature of 4° C. (the number of cycles is optimally 30–50 to determine the appearance or lack of a PCR product and optimally 8–30 cycles if an estimation of the starting quantity of cDNA from the RT reaction is to be made); 10 microlitre aliquots are then run out on 1% 1×TBE gels stained with ethidium bromide with PCR product, if present, sizes estimated by comparison to a 100 bp DNA Ladder (Gibco BRL, Life Technologies). Alternatively if the PCR products are conveniently labelled by the use of a labelled PCR primer (e.g. labelled at the 5'end with a dye) a suitable aliquot of the PCR product is run out on a polyacrylamide sequencing gel and its presence and quantity detected using a suitable gel scanning system (e.g. ABI Prism™ 377 Sequencer using GeneScan™ software as supplied by Perkin Elmer).

RT/PCR controls may include +/− reverse transcriptase reactions, 16s rRNA primers or DNA specific primer pairs designed to produce PCR products from non-transcribed *Staphylococcus aureus* WCUH29 genomic sequences.

To test the efficiency of the primer pairs they are used in DNA PCR with *Staphylococcus aureus* WCUH29 total DNA. PCR reactions are set up and run as described above using approx. 1 microgram of DNA in place of the cDNA and 35 cycles of PCR.

Primer pairs which fail to give the predicted sized product in either DNA PCR or RT/PCR are PCR failures and as such are uninformative. Of those which give the correct size product with DNA PCR two classes are distinguished in RT/PCR: 1. Genes which are not transcribed in vivo reproducibly fail to give a product in RT/PCR; and 2. Genes which are transcribed in vivo reproducibly give the correct size product in RT/PCR and show a stronger signal in the +RT samples than the signal (if at all present) in −RT controls.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 5

<210> SEQ ID NO 1
<211> LENGTH: 2106
<212> TYPE: DNA
<213> ORGANISM: Staphylococcus aureus

<400> SEQUENCE: 1

atgaataacg gttttttcaa tagcgacttt gattcaattt ttcgaagaat gatgcaagat 60

atgcaaggtt caaatcaagt cggtaacaaa aagtactata ttaatggtaa agaagtttca 120

cctgaagaac tagcgcaact cacacaacaa ggtagcaatc aatctgctga acaaagtgcg 180

caagcttttc aacaagcagc acaaagacaa caagggcaac aaggtggcaa cggcaattat 240

ttagaacaaa ttggtcgtaa ccttacgcaa gaagcacgtg acggtttatt agatccagtc 300

attggtcgtg ataaagaaat tcaagaaact gctgaagtct taagtagacg aactaaaaac 360

aatcctatat tagttggaga agctggtgtt ggtaaaactg cgattgttga aggtttagca 420

caggcaatcg ttgaaggaaa tgtaccagca gcaatcaaag acaaagaaat tatttctgta 480

gatatttcat cattagaagc tggaacgcaa tatcgtggtg cttttgaaga aaatattcaa 540

aaattaatcg aaggtgttaa atcttcacaa aatgccgtac tattctttga tgaaatccat 600

caaattatcg gttcaggtgc cacaggaagt gattcaggta gcaaagggtt atctgatatt 660

ttgaaacctg cattaagtcg tggtgagatt tctattattg gtgcaacaac acaagatgaa 720

tatcgaaaca atattcttaa agatgctgca ttaacgcgca gatttaatga agtgcttgtt 780

aatgaaccaa gcgctaaaga tactgttgaa attttaaaag gtattcgcga aaaattcgaa 840

gaacaccatc aagtaaaatt accagatgac gtattaaaag catgtgttga cttatcaatt 900

caatatattc cacaacgatt attaccagat aaagcaatcg atgtgttaga tattacagca 960

gcacatttat ctgcgcaaag tccagctgtc gataaagttg aaactgaaaa acgaatttct 1020

gaattagaaa atgataaacg taaagcagta agtgctgaag aatataaaaa agctgacgac 1080

attcaaaatg aaatcaaatc attacaagat aaattagaaa acagtaatgg tgaacatact 1140

gctgttgcta cagttcatga tatttcagat actattcaac gattaactgg cattccagtt 1200

tctcaaatgg atgataacga tattgaacgt ttaaaaaata tttctaatcg tttaagaagt 1260

aaaatcatag gtcaagatca agctgtagaa atggtttcac gtgcaattcg ccgtaatcgc 1320

-continued

gctgggtttg atgacggcaa ccgtccaatt ggtagtttct tatttgtcgg ccctactggt 1380

gttggtaaaa cagagcttgc taaacaatta gcaattgatc tatttggtaa taaagatgca 1440

cttattcgac ttgatatgag tgaatatagt gacacaacag ctgtttcaaa aatgattggt 1500

acaactgctg gttatgtcgg ttatgatgac aattcaaata cgttaactga aaaagtacgc 1560

cgtaatccat actcagtcat tctatttgat gaaatcgaaa aagcaaatcc acaaatttta 1620

acattgttat tacaagtaat ggatgatggt aatttgactg atggtcaagg taatgtcatc 1680

aactttaaaa atacaattat tatttgtaca tcaaatgctg gctttggcaa tggcaatgac 1740

gctgaagaaa aagatattat gcacgaaatg aaaaaattct tccgccctga attccttaac 1800

cgcttcaacg gcatcgttga attcttacat ttagataaag atgcattgca agatatcgtc 1860

aacttattat tagacgatgt acaagttaca ttagataaaa aaggtattac gatggacgtt 1920

tctcaagatg cgaaagattg gttaattgaa gaaggctatg atgaagaatt aggtgcacgt 1980

ccattaagac gtattgttga acagcaagta cgtgacaaaa ttacagatta ttatctagat 2040

catacagacg ttaaacatgt ggatatagat gttgaggata acgaattagt cgtaaaaggt 2100

aaataa 2106

<210> SEQ ID NO 2
<211> LENGTH: 701
<212> TYPE: PRT
<213> ORGANISM: Staphylococcus aureus

<400> SEQUENCE: 2

Met Asn Asn Gly Phe Phe Asn Ser Asp Phe Asp Ser Ile Phe Arg Arg
 1   5   10   15

Met Met Gln Asp Met Gln Gly Ser Asn Gln Val Gly Asn Lys Lys Tyr
            20   25   30

Tyr Ile Asn Gly Lys Glu Val Ser Pro Glu Glu Leu Ala Gln Leu Thr
        35   40   45

Gln Gln Gly Ser Asn Gln Ser Ala Glu Gln Ser Ala Gln Ala Phe Gln
    50   55   60

Gln Ala Ala Gln Arg Gln Gln Gly Gln Gln Gly Gly Asn Gly Asn Tyr
65   70   75   80

Leu Glu Gln Ile Gly Arg Asn Leu Thr Gln Glu Ala Arg Asp Gly Leu
                85   90   95

Leu Asp Pro Val Ile Gly Arg Asp Lys Glu Ile Gln Glu Thr Ala Glu
            100   105   110

Val Leu Ser Arg Arg Thr Lys Asn Asn Pro Ile Leu Val Gly Glu Ala
        115   120   125

Gly Val Gly Lys Thr Ala Ile Val Glu Gly Leu Ala Gln Ala Ile Val
    130   135   140

Glu Gly Asn Val Pro Ala Ala Ile Lys Asp Lys Glu Ile Ile Ser Val
145   150   155   160

Asp Ile Ser Ser Leu Glu Ala Gly Thr Gln Tyr Arg Gly Ala Phe Glu
                165   170   175

Glu Asn Ile Gln Lys Leu Ile Glu Gly Val Lys Ser Ser Gln Asn Ala
            180   185   190

Val Leu Phe Phe Asp Glu Ile His Gln Ile Ile Gly Ser Gly Ala Thr
        195   200   205

Gly Ser Asp Ser Gly Ser Lys Gly Leu Ser Asp Ile Leu Lys Pro Ala
    210   215   220

-continued

```
Leu Ser Arg Gly Glu Ile Ser Ile Ile Gly Ala Thr Thr Gln Asp Glu
225                 230                 235                 240

Tyr Arg Asn Asn Ile Leu Lys Asp Ala Ala Leu Thr Arg Arg Phe Asn
                245                 250                 255

Glu Val Leu Val Asn Glu Pro Ser Ala Lys Asp Thr Val Glu Ile Leu
            260                 265                 270

Lys Gly Ile Arg Glu Lys Phe Glu Glu His His Gln Val Lys Leu Pro
        275                 280                 285

Asp Asp Val Leu Lys Ala Cys Val Asp Leu Ser Ile Gln Tyr Ile Pro
    290                 295                 300

Gln Arg Leu Leu Pro Asp Lys Ala Ile Asp Val Leu Asp Ile Thr Ala
305                 310                 315                 320

Ala His Leu Ser Ala Gln Ser Pro Ala Val Asp Lys Val Glu Thr Glu
                325                 330                 335

Lys Arg Ile Ser Glu Leu Glu Asn Asp Lys Arg Lys Ala Val Ser Ala
            340                 345                 350

Glu Glu Tyr Lys Lys Ala Asp Asp Ile Gln Asn Glu Ile Lys Ser Leu
        355                 360                 365

Gln Asp Lys Leu Glu Asn Ser Asn Gly Glu His Thr Ala Val Ala Thr
    370                 375                 380

Val His Asp Ile Ser Asp Thr Ile Gln Arg Leu Thr Gly Ile Pro Val
385                 390                 395                 400

Ser Gln Met Asp Asp Asn Asp Ile Glu Arg Leu Lys Asn Ile Ser Asn
                405                 410                 415

Arg Leu Arg Ser Lys Ile Ile Gly Gln Asp Gln Ala Val Glu Met Val
            420                 425                 430

Ser Arg Ala Ile Arg Arg Asn Arg Ala Gly Phe Asp Asp Gly Asn Arg
        435                 440                 445

Pro Ile Gly Ser Phe Leu Phe Val Gly Pro Thr Gly Val Gly Lys Thr
    450                 455                 460

Glu Leu Ala Lys Gln Leu Ala Ile Asp Leu Phe Gly Asn Lys Asp Ala
465                 470                 475                 480

Leu Ile Arg Leu Asp Met Ser Glu Tyr Ser Asp Thr Thr Ala Val Ser
                485                 490                 495

Lys Met Ile Gly Thr Thr Ala Gly Tyr Val Gly Tyr Asp Asp Asn Ser
            500                 505                 510

Asn Thr Leu Thr Glu Lys Val Arg Arg Asn Pro Tyr Ser Val Ile Leu
        515                 520                 525

Phe Asp Glu Ile Glu Lys Ala Asn Pro Gln Ile Leu Thr Leu Leu Leu
    530                 535                 540

Gln Val Met Asp Asp Gly Asn Leu Thr Asp Gly Gln Gly Asn Val Ile
545                 550                 555                 560

Asn Phe Lys Asn Thr Ile Ile Ile Cys Thr Ser Asn Ala Gly Phe Gly
                565                 570                 575

Asn Gly Asn Asp Ala Glu Glu Lys Asp Ile Met His Glu Met Lys Lys
            580                 585                 590

Phe Phe Arg Pro Glu Phe Leu Asn Arg Phe Asn Gly Ile Val Glu Phe
        595                 600                 605

Leu His Leu Asp Lys Asp Ala Leu Gln Asp Ile Val Asn Leu Leu Leu
    610                 615                 620

Asp Asp Val Gln Val Thr Leu Asp Lys Lys Gly Ile Thr Met Asp Val
625                 630                 635                 640

Ser Gln Asp Ala Lys Asp Trp Leu Ile Glu Glu Gly Tyr Asp Glu Glu
```

-continued

```
                645                 650                 655

Leu Gly Ala Arg Pro Leu Arg Arg Ile Val Glu Gln Gln Val Arg Asp
            660                 665                 670

Lys Ile Thr Asp Tyr Tyr Leu Asp His Thr Asp Val Lys His Val Asp
        675                 680                 685

Ile Asp Val Glu Asp Asn Glu Leu Val Val Lys Gly Lys
    690                 695                 700


<210> SEQ ID NO 3
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Staphylococcus aureus

<400> SEQUENCE: 3

atgaataacg gttttttcaa tagc                                          24


<210> SEQ ID NO 4
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Staphylococcus aureus

<400> SEQUENCE: 4

ttatttacct tttacgacta attc                                          24


<210> SEQ ID NO 5
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Staphylococcus aureus

<400> SEQUENCE: 5

gctcctaaaa ggttactcca ccggc                                         25
```

What is claimed is:

1. An isolated polypeptide comprising SEQ ID NO:2.

2. A composition comprising the isolated polypeptide of claim **1** and a carrier.

3. The isolated polypeptide of claim **1**, wherein the isolated polypeptide comprises a heterologous amino acid sequence fused to SEQ ID NO:2.

4. A composition comprising the isolated polypeptide of claim **3** and a carrier.

5. The isolated polypeptide of claim **1**, wherein the isolated polypeptide consists of SEQ ID NO:2.

6. A composition comprising the isolated polypeptide of claim **5** and a carrier.

* * * * *